DOCUMENT RESUME

ED 476 857 TM 034 950

AUTHOR	Henninger, Michael; Balk, Michael
TITLE	With or without? The Role of Reference Values in Evaluation Feedback to Teachers.
PUB DATE	2003-04-00
NOTE	17p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	*Evaluation Utilization; *Expectation; *Feedback; Foreign Countries; *Goal Orientation; Medical Education; *Medical School Faculty; Medical Students; *Student Evaluation of Teacher Performance
IDENTIFIERS	Germany

ABSTRACT

This study was conducted to see if evaluative feedback from students that includes reference values (students' expectations on teaching quality) has an effect on acceptance by teachers and relevance for them. Also examined was the degree to which teachers' goal setting is influenced by feedback of student ratings with or without reference values. The sample consisted of 35 teachers at a German faculty of medicine. The 10 in the experimental group received personal feedback of evaluation results and students expectancies as reference values. Experimental group 2 (n-11) received personal feedback without reference values, and the third group did not receive any feedback. Findings show that including reference values does not result in higher acceptance of evaluation feedback. However, in this study, acceptance of evaluation results was so high overall that a ceiling effect may have occurred. Teachers in both experimental groups significantly enhanced their goal setting after the personal feedback, while teachers in the comparison group did not alter their goal setting. Findings are interpreted as supporting personal feedback to teachers. (Contains 38 references.) (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 476 857

84th Annual Meeting of the American Educational Research Association
"Accountability for Educational Quality. Shared Responsibility"
April 21 to 25, 2003
Chicago, Illinois

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

M. Henninger

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

With or without?
The role of reference values in evaluation feedback to teachers

Michael Henninger & Michael Balk
University of Regensburg, Germany

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

The problem in focus

One of the main problems in improving the educational quality of university teaching is how to use student ratings' feedback which is provided to teachers (Henninger, 2002; Marsh & Roche, 1993, 1997; McKeachie, 1997). In order to inform teachers about strengths and weaknesses of teaching quality, many different forms of student ratings' feedback have been developed and investigated (Centra, 1973; Cohen, 1980; Marsh & Roche, 1993; McKeachie et al., 1980; Miller, 1971; Murray & Smith, 1989; Wilson, 1986). But unfortunately, meta-studies document only a moderate positive effect of student ratings' feedback to teachers (Cohen, 1980; L'Hommedieu, Menges & Brinko, 1990). Therefore, still the question arises of how to design feedback to teachers which is effective for a change in teaching actions.

Drawing on feedback and action theory (Frese & Zapf, 1994; Kluger & DeNisi, 1996), feedback information has to be actively processed in order to have an effect on cognitive-motivational parameters, e.g. teachers' goal setting (Locke, 2001). But, teachers only actively process feedback when it is accepted by and relevant for them. Thus, we argue that it has to be examined thoroughly how to affect and enhance teachers' acceptance and relevance of the feedback (Henninger, 2002). Surprisingly, this question has not been investigated in depth.

In german speaking countries, we find that teachers' acceptance and their perceived relevance of using student ratings as informative feedback frequently is very low (Henninger, 2001; Henninger & Balk, 2001, 2002; Spiel, Wolf & Popper, 2002). Many teachers doubt that student ratings are valid information (Rindermann, 2002). They frequently argue that ratings may be biased because of high and unrealistic expectations of the students. Astonishingly one of the solutions for this problem – assessment of expectations of students - are seldom included in the feedback as

TM034950

BEST COPY AVAILABLE

reference values for the evaluative ratings (Cohen, 1980; Henninger & Balk, 2002; Henninger, Balk & Mandl, 1998). Besides the huge research on the validity of student ratings (d'Appollonia & Abrami, 1997; Greewald, 1997; Greenwald & Gillmore, 1997; Marsh & Roche, 1997, 1999), there is not much research which examines the effect of reference values such as students' expectancies on teachers' acceptance and relevance of evaluative feedback. We assume that integrating reference values into the feedback, such as students' expectations, would enhance the transparency for the teachers and their trust in student ratings. Given a sufficient level of acceptance and perceived relevance for their teaching, feedback effects on cognitive-motivational variables of the action process, especially on teachers' goal setting, can be investigated more differentiatedly.

Empirical findings – state of the art

Two meta-studies documented effect sizes of student ratings' feedback which are moderately positive (see Cohen, 1980; L'Hommedieu et al., 1990). Cohen (1980) reviewed 17 studies with 22 comparisons of feedback effects with control groups. The average effect size was *d*=.38 (Cohen, 1977). Feedback was given at midterm - mainly printouts of means, standard deviations and frequencies were sent to the teachers - and the major dependent variable was course evaluation at the end of the term, besides students' attitudes and achievement measures in some studies. Improvements were found especially concerning specific evaluation dimensions: instructor's skill (*d*=.47) and feedback to students (*d*=.40). Other students' measures showed heterogenous effects. Teaching experience, reference values such as teachers' self evaluation or comparisons to other teachers' values as well as the time to realize changes in teaching did not influence feedback effects. When feedback was delivered by an evaluator or within a counseling intervention, higher effects were found: general evaluation *d*=.64, instructor's skill *d*=.59, feedback to students *d*=.50. It seems that especially dimensions of teaching quality which actively can be influenced by teachers' actions improved after feedback. Other dimensions which are affected by content or course size did not improve.

L'Hommedieu et al. (1990) raised the question why the empirical evidence to use the feedback of student ratings for improvement of teaching quality is not more convincing. The authors meta-analyzed 28 studies which were selected with familiar criteria as in Cohen's (1980) review (e.g. control group design, midterm feedback, endterm student ratings as dependent variable). They also found a moderate, positive effect (d=.342; Glass, McGaw & Smith, 1981). They conclude: "Certainly this degree of improvement is not sufficient to justify the expense of designing, administering, and analyzing student rating instruments. Yet it seems inconceivable that the opinions of a professor's students can mean so little" (L'Hommedieu et al., 1990, p. 233). Therefore, they profoundly analyzed characteristics of feedback

research in the domain of teaching evaluation - especially threats of validity (Cook & Campbell, 1979). They come to the conclusion that within field settings it is not easy to control these methodological problems which tend to attenuate feedback effects. We summarize their recommendations for future research which to us seem especially important: (1) Random assignment of teachers to experimental and control groups in order to enhance internal validity. (2) Use end-term feedback to enhance construct validity, because this is applied in practice much more frequently than midterm feedback. (3) Control teachers' attention to the feedback. (4) Assess known moderating variables, such as teachers' perceptions of and reactions to the feedback. (5) Control losses of subjects (teachers, students).

All in all, looking at the research literature on student ratings feedback to teachers it can be stated that the effects frequently are lower than expected. But, authors of both studies argue that research on student ratings' feedback still is an important issue in education and educational psychology in order to foster the effectiveness of feedback and gain insights in the effects of variables that still are poorly understood. It is suprisingly that research on this topic had decreased since the 1990ies. Many researchers state that one should focus on investigating counseling feedback (Marsh & Roche, 1993, 1997; Rindermann, 2001; Rindermann & Kohler, 2003). We do not fully agree with that, because from an action theoretical point of view, key variables such as teachers' goal setting, acceptance of feedback, the role of self-efficacy and the relevance of the feedback for teachers' further planning of their teaching still have not been examined sufficiently. We think that using the concept of action regulation can help to clarify which aspects serve as moderator variables, such as teachers' reactions to feedback as argumented by L'Hommedieu et al. (1990; see also Rotem & Glasman, 1979). Before we elaborate this theoretical avenue for investigating feedback effects in more detail, we like to point out our major critique of past research - agreeing with the recommendations of L'Hommedieu et al. (1990).

(1) Research focussed on mid-term feedback. In practice, end-term feedback is used much more frequently for improving teaching quality. Furthermore, reference values such as self evaluation or evaluation scores of other teachers did not show much effect (Cohen, 1980). Reference values such as students' expectations on teaching quality have not been researched as a treatment variation yet.

(2) Attention payed to the feedback and active processing of the information as necessary prerequisites for modification of actions (Ilgen & Davis, 2000; Kluger & DeNisi, 1996) only can be controlled by using personal delivered feedback. But many studies used feedback printouts which were sent to the teachers.

(3) Little is known about teachers' reactions to and perceptions of the feedback. A variety of potentially moderating variables can be operationalized by applying an action theoretical framework, such as acceptance and relevance of the feedback

for the teacher or self-efficacy - especially in the context of teachers' distrust in teaching evaluation as outlined at the beginning of this paper.

(4) The dependent variable which is frequently assessed in research on student ratings feedback is the evaluation of the course after feedback has been provided to teachers (L'Hommedieu et al., 1990). Cognitive-motivational variables which are closer to the person who is taking action in the course seldom are investigated, such as teachers' goal setting.

Theoretical framework of the study

Applying the action process according to Frese & Zapf (1994; Henninger, 2000) investigating feedback effects in the domain of teaching evaluation, it is shown in figure 1 that feedback is linked with action goals. This is the dependent cognitive-motivational variable which is nearest to providing feedback to the person. There are several action steps which are inbetween the persons' regulation process before the execution of the next actions. This means that assessing feedback effects by student ratings after the feedback has been given to the teacher - as in the studies documented above - neglects at least to investigate the effect on teachers' action goals. If teachers do not change their goal setting, they would also not change their teaching actions. Therefore, we suggest goal setting (Balk, 2000; Frese & Zapf, 1994; Henninger, 2000; Henninger & Balk, 2002; Ilgen & Davis, 2000; Locke, 2001; Locke & Latham, 1990) as a central, theoretically derived dependent variable as a direct person related indicator for effects of evaluation feedback.

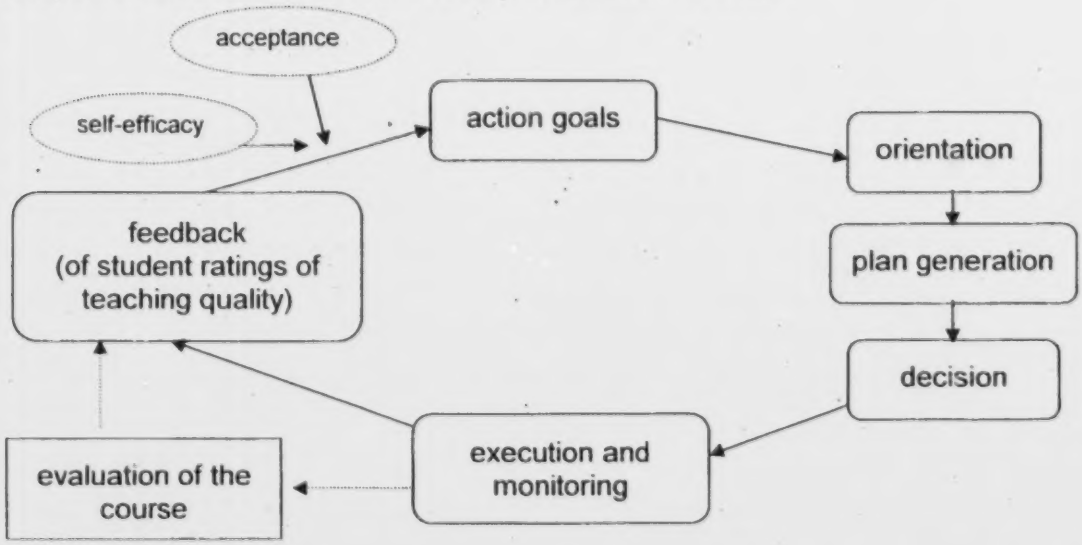

Figure 1: Action process following Frese und Zapf (1994; see Henninger, 2000)

BEST COPY AVAILABLE

In the research literature on action regulation it is discussed which variables influence the relationship between feedback and goal setting. First of all, feedback must be accepted by and relevant for the teacher in order to be actively processed and used for further goal setting processes (Henninger & Balk, 2001; Kluger & DeNisi, 1996; L'Hommedieu et al., 1990; Nease, Mudgett & Quinones, 1999; Rotem & Glasman, 1979). When a sufficient level of acceptance and relevance is given, it is not sure if the teacher is convinced to be able to realize changes in teaching behavior when necessary. Maybe the teacher wants to increase goal setting, i.e. to enhance the structure of a lecture, but has low self-efficacy on this evaluation dimension. As a consequence, he/she may decide not to increase goal setting because confidence in own abilities is low. Therefore, self-efficacy (Balk, 2000; Bandura, 1997; Cohen, 1980; Nease et al., 1999; Young & Kline, 1996) has to be assessed as a variable which is plausible to influence goal setting.

Goals of the study

As argued above, variables which are prone to influence the active processing of the feedback information are teachers' acceptance and perceived relevance of the feedback. Thus, we want to find out if evaluative feedback which includes reference values, i.e. students' expectations on teaching quality, has an effect on acceptance and relevance. This reference value still has not been clarified by past research.

Moreover, given a sufficiently degree of acceptance and relevance, we want to find out which effect student ratings' feedback has on cognitive-motivational variables relevant for teachers' action process (Henninger, 2002; Henninger & Balk, 2002). As one important aspect of action regulation, we want to figure out inhowfar teachers' goal setting is influenced by this feedback. Unfortunately, research on feedback effects in the domain of university teaching still has put not very much attention on this theoretically very important dependent variable. As has been shown in broad research in contexts different from evaluation at university, goal setting is a major contributor for motivation and high quality performance (see e.g. Locke, 2001; Locke & Latham, 1990). But, teachers' self-efficacy may hinder or foster goal setting - the degree to which they can influence evaluation dimensions (Bandura, 1997; Cohen, 1980). Thus, we also assessed this possibly confounding variable.

Research questions:

(1) *What effect does student ratings' feedback including reference values - students' expectations - have on teachers' acceptance and relevance for further planning?*

(2) *Inhowfar is teachers' goal setting influenced by personal feedback of student ratings, with and without reference values?*

Methods

The sample consists of *N*=35 university teachers at the Faculty of Medicine in Regensburg, Germany (8 full professors, 7 assistant professors and 20 medical doctors), who participated at the study on a voluntarily basis. We established a three group experimental design by randomly assigning 21 teachers to the experimental groups (subjects were drawn by lot by two researchers): experimental group 1 (*n*=10) received personal feedback of evaluation results, i.e. the evaluative judgements and students' expectancies as reference values (see figure 2). Experimental group 2 (*n*=11) received personal feedback of evaluation results, i.e. only the evaluative judgements. Student ratings were collected at the end of the course and feedback was presented when the semester was over (end-term feedback). The contrast group (*n*=14) did not receive any feedback. As Wilkinson and the Task Force on Statistical Inference (1999) argue, the term control group should be used very cautiously. Therefore, they suggest the term contrast group. In our case, the subjects in the contrast group were teachers who conducted a course in the semester when the evaluation took place for the teachers of the experimental group. Furthermore, as the teachers in the experimental groups, the teachers in the contrast group maintained their teaching in the following semester. This is important to note, because not all faculty members are involved in courses in each semester, concentrating e.g. periodically on research or clinical practice.

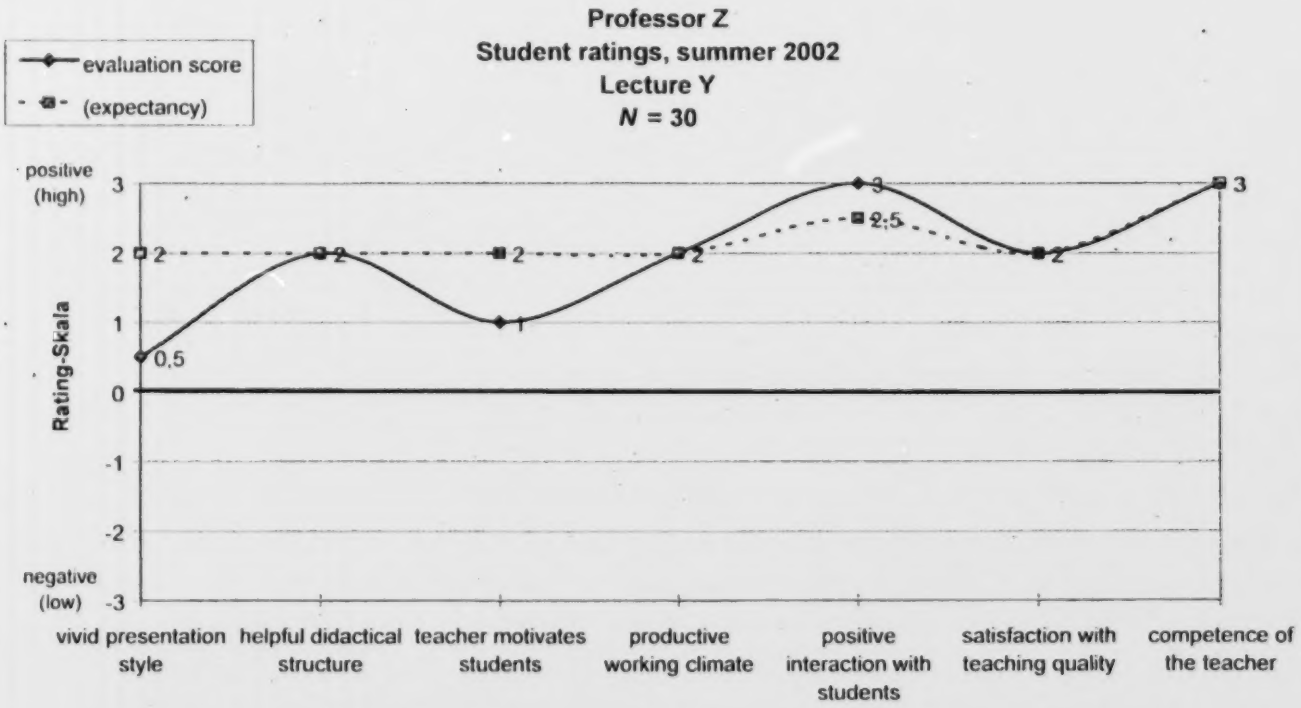

Figure 2: Grafically displayed feedback of evaluation scores augmented with reference values (students' expectations on teaching quality)

The treatment is a personal feedback of evaluation results. We decided to personally deliver the feedback in order to control for teachers' attention and processing of the information (see Cohen, 1980; Kluger & DeNisi, 1996; L'Hommedieu et al., 1990). Moreover, 'personal' means that the results are only presented to the teacher, not to faculty members or the public, in order to establish a trustful cooperation which we view as necessary for teacher's acceptance of the evaluation. Presentation of feedback is standardized and video-taped to control for confounding variables such as reactivity of the evaluator. The evaluator has a one-to-one 10-minute meeting with the individual teacher. The results are grafically visualized[1] and delivered by verbal descriptions of the evaluator in order to generate and guide teachers' attention to each aspect of the feedback information. The descriptions follow a specific sequence and point to the graphics, the values (medians) and their semantic meaning, e.g. that the higher evaluative judgements compared to the expectancies is positive in case of experimental group 1 who additionally received reference values. With this treatment, teachers' active processing of the information should be facilitated. We decided to use medians as the more robust measure compared to the arithmetical mean, because there were several smaller courses with less than 30 students. This selection of statistics was also explained to the teachers. In addition to the graphics, teachers' received a table with medians for three items which assessed students' interest and evaluations concerning the room and time of the course.

As dependent variables which are prerequisites for integrating feedback information in teachers' action process, we assessed teachers' acceptance ("It is good to receive a personal feedback of evaluation results."; "The graphics are clearly displayed."; "I think it's good to graphically display the evaluation results.") and subjective relevance of the feedback ("The evaluation results are relevant for further planning of teaching."). Teachers judged their agreement on each item on 7-point rating-scales (6=agree; 0=disagree).

As the dependent cognitive-motivational variable of the action regulation process, we assessed teachers' goal setting for each evaluation dimension (e.g. "Considering the following term, my goal is to make the didactic structure more helpful.") with a questionnaire using 7-point rating-scales (6=agree; 0=disagree). Furthermore, we assessed teachers' self-efficacy on each evaluation dimension (e.g. "I trust in my ability to make the didactic structure more helful."; 7-point rating-scales; 6=high self-efficacy, 0=low self-efficacy). Then, we calculated a goal setting index which included only aspects with the median of self-efficacy scores equal and higher *Md*=4 in order to assure that teachers can influence the evaluation dimensions (Cohen, 1980). The intercorrelations between these self-efficacy items are highly positive for four out of

[1] Please note that the graphics are a visualization for the teachers in order to help them to process the evaluation data.

six items, ranging from r=.61 to r=.76 (p<.001). The intercorrelations of two items are lower and more inconsistently (r=.34 to r=.71; p<05). With this selection procedure, the following evaluation dimensions were included in the goal setting index: vivid presentation style, helpful didactic structure, motivating students, productive working climate, positive interaction with students, high quality of the course, and high interest of the students. The intercorrelations between these dimensions included in the goal setting index are highly positive, ranging from r=.81 to r=.95 (p<.001).

Teachers filled in questionnaire 1 to assess the baseline for goal setting and self-efficacy prior to the feedback. The questionnaire was sent via e-mail, fax or by mail according to the preferences of the teacher and sent back before treatment. Nine subjects of the experimental groups returned the questionnaire in person at the meeting when the feedback was provided. Questionnaire 2 was administered directly after the feedback for the experimental groups and assessed acceptance and relevance of the feedback, and goal setting. The contrast group received both questionnaires via e-mail. Questionnaire 2 for the contrast group assessed goal setting. The time between t1 and t2 was considered to be comparable with the experimental group, but due to organizational constraints at the faculty, that was not possible in any case. Unfortunately, the contrast group suffered from a severe drop-out. Five subjects returned questionnaire 2, instead of several reminders for all subjects in the contrast group.

We first of all conducted - due to the small sample size - nonparametrical data analyses. To examine differences between groups, we conducted Mann-Whitney U-tests. To investigate changes within groups from t1 to t2, we conducted Wilcoxon-tests.

Results

With regard to research question 1, we investigated the effect of student ratings' feedback including reference values, i.e. students' expectations, on teachers' acceptance and relevance for further planning compared with mere feedback of evaluative judgements. Table 1 shows the results of the Mann-Whitney U-test.

In both groups, means of all three items which assess teachers' acceptance are very high. The means lie inbetween the values 5 and 6 of the rating scales (6=high; 0=low). Only one comparison is statistically significant. The mean of the item "I think it's good to graphically display the evaluation results" is significantly lower for the group who received feedback including reference values (p<.05).

The feedback's relevance for teachers' further planning of course is also very high. In both groups, the means lie inbetween the values 5 and 6 of the rating scales (6=high; 0=low). Means for both groups do not differ significantly.

Table 1

Group differences for teachers' acceptance and relevance of the feedback between groups who received student ratings' feedback with and without reference values

	Reference values		No reference values		
	M	*SD*	*M*	*SD*	*Z*
Acceptance: "good to receive feedback"	5.90	0.32	6.00	0.00	-1.05
Acceptance: "graphics are clearly displayed"	5.67	0.50	5.73	0.65	-0.60
Acceptance: "good to graphically display feedback"	5.10	1.20	5.91	0.30	-2.09**
Relevance: "results are relevant"	5.30	0.82	5.18	1.08	-0.08

**$p < .05$ (one-sided).

Research question 2 investigates the influence of personal feedback with and without reference values on teachers' goal setting. Table 2 shows the results of the Wilcoxon test comparing within groups.

Table 2

Within-group comparisons for teachers' goal setting

	t1		t2		
Treatment	*M*	*SD*	*M*	*SD*	*Z*
Reference values	1.93	2.32	3.20	1.95	-1.69**
No reference values	2.21	2.14	4.21	1.26	-2.02**
No feedback	1.80	2.19	2.00	2.41	-1.34

**$p < .05$ (one-sided).

In both experimental groups, the means of goal setting significantly are higher after the feedback intervention ($p<.05$). After the personal feedback, the means lie above the rating scale mid-point 3. This significant increase of the means from t1 to t2

indicates an enhancement in goal setting regarding the evaluation dimensions, e.g. teachers intend to make the didactic structure more helpful. In the contrast group, the means at t1 and t2 lie below the rating-scale mid-point 3.

In table 3, 4 and 5, the results of the Mann-Whitney U-tests for between-groups comparisons are shown. Due to missing values, the sample sizes are n=8 for experimental group 1 ('reference values'), n=9 for experimental group 2 ('no reference values') and n=5 for the contrast group ('no feedback').

Table 3

Between-groups' comparisons for teachers' goal setting

	Reference values		No reference values		
	M	*SD*	*M*	*SD*	*Z*
t1	1.93	2.32	2.21	2.14	-0.30
t2	3.27	1.94	4.18	1.33	-1.16

The results of table 3 indicate for t1 and t2 that means of goal setting are not different for the group who received student ratings' feedback including reference values and for the group who was provided with feedback of evaluative judgements without reference values.

Table 4

Between-groups' comparisons for teachers' goal setting

	Reference values		No feedback		
	M	*SD*	*M*	*SD*	*Z*
t1	1.93	2.32	1.80	2.49	-0.08
t2	3.27	1.94	2.00	2.41	-0.88

There is no significant difference at t1 as well as t2 between the experimental group who received student ratings' feedback including reference values and the contrast group who did not receive any feedback (table 4).

Table 5

Between-groups' comparisons for teachers' goal setting

	No reference values		No feedback		
	M	*SD*	*M*	*SD*	*Z*
t1	2.21	2.14	1.80	2.49	-0.42
t2	4.18	1.33	2.00	2.41	-1.74**

**$p < .05$ (one-sided).

Table 5 shows that means in goal setting at t1 are not different between the experimental group who received feedback without reference values and the contrast group. At t2 there is a significant difference between the groups ($p<.05$): The experimental group's mean of goal setting is significantly higher.

Discussion

Concerning research question 1, results show that including reference values does not lead to a higher acceptance of evaluation feedback as compared with the group without reference values. In contrast, feedback without reference values produces a higher effect on one item of acceptance, i.e. teachers' judgement that it is good to graphically display the results. At first sight, both results - the general and the specific one – do not confirm the theoretically-based assumption that including reference values leads to a higher acceptance of feedback results. However, when taking a second look, one can see that both parameter indicate a high level of acceptance of the evaluation results and a very small standard deviation. It seems that we have a ceiling effect, thus a variation could not create any difference. Interestingly, the high acceptance and relevance in this study are in accordance with a large-sample survey which was conducted across all faculties at the University of Regensburg (Henninger & Fuchs, 2002). In the latter study, teachers also indicated that student ratings are valid, accepted information for teachers and are relevant for them.

Thus, we can state that the teachers' acceptance of the personal feedback and the relevance for their further planning is very high in the study presented here. With our intervention – personally-delivered feedback which only is provided to the individual teacher - we reached a sufficiently high level of these two variables. As argued in the beginning of our paper, perceived acceptance and relevance of the feedback are crucial for further effects of the evaluative information within the action process of teachers.

Given teachers' acceptance of the feedback and it's relevance, we investigated the effects of personal feedback on teachers' action goals in research question 2. Results let us state that it is possible to document changes in teachers' goal setting. This supports our argumentation at the beginning of this paper to investigate feedback effects on this dependent cognitive-motivational variable. Teachers in both experimental groups significantly enhanced their goal setting after the personal feedback. Teachers in the contrast group did not change their goal setting. We propose that by personal delivering the feedback to the teachers in the experimental groups an effect of the specific treatments - feedback with and without reference values - possibly could be covered (see also Marsh & Roche, 1993; Wilson, 1986).

The comparisons between groups show an interesting pattern supporting this proposition: There is no difference between both experimental groups. This means that including reference values did not strengthen the effect of feedback on goal setting in this study. The non-significant difference in goal setting between experimental group 1 - feedback with reference values - and the contrast group is surprisingly because there is no increase in this parameter from t1 to t2 in the contrast but in the experimental group 1. This possibly can be explained with the small sample size in the contrast group which may tendencially be a positive selection of teachers. Only teachers who are motivated to a certain degree to participate in an educational study may return the questionnaire 2, especially under the organizational constraints at the medical faculty. This may influence their judgements on goal setting which are slightly higher than at t1 for all groups, thus a difference between experimental group 1 and the contrast group may masked. In a replication study, we want to look at that point again by realizing higher sample sizes in all three groups. Furthermore, it is interesting that teachers in the experimental group who received traditional feedback, i.e. without reference values, enhanced their goal setting most. This is shown by the within-group comparison as well as by comparison with the contrast group.

All in all, we interpret the findings to be a positive support for personal feedback to teachers. Interestingly, unlike our assumption, reference values seem not to play a central role for teachers' acceptance and relevance of the feedback. It worked out to create a feedback intervention in this study which is highly accepted by teachers, with or without reference values. Therefore, results of this study are in line with the consequences drawn from feedback research in the domain of university teaching by L'Hommedieu et al. (1990) and Cohen (1980): Feedback has a positive effect on teaching actions, but frequently is attenuated due to methodological problems. We tried to control central methodological problems in this study, such as random assignment, using end-term feedback, control attention of teachers to the feedback by personally delivering it and assessing teachers' reactions to the feedback, i.e. acceptance and relevance. Unfortunately, in this study drop out especially in the

contrast group was a severe problem. As mentioned above, we conduct a replication study with larger sample size. Furthermore, we assess general and specific goal setting (Locke, 2001; Locke & Latham, 1990) in order to further differentiate the dependent cognitive-motivational variable. Moreover, we assess motivational variables according to Deci & Ryan (1993; Ryan & Deci, 2000) that may influence goal setting and may be confounding variables for feedback effects as discussed above concerning the possibly positive selection of teachers at t2 in the contrast group. Of course, to investigate feedback effects in field settings makes it difficult to control possible confounding variables such as selection effects, positivity of evaluation scores etc. (Cook & Campbell, 1979). Thus, it makes sense from our point of view to conduct experimental studies in order to control certain variables more strongly.

Finally, one may argue that the improvement of evaluation scores after the feedback should be one central dependent variable for feedback effects. Therefore, in future studies the link between feedback, teachers' goal setting and the execution of the action in the next course is of high valence. But as we would suggest, student ratings of teachers actions in the course are not the only measures for feedback effects. We would recommend to additionally use expert ratings of teachers actions by focusing on goal-related behavior, and e.g. teachers' learning achievement in instructional training courses.

References

Balk, M. (2000). *Evaluation von Lehrveranstaltungen. Die Wirkung von Evaluationsrückmeldung* [Evaluation of courses. The effect of evaluation feedback]. Frankfurt am Main: Lang.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: Freeman.

Centra, J. A. (1973). Effectiveness of student feedback in modifying college instruction. *Journal of Educational Psychology, 65*, 395-401.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). New York: Academic.

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education, 13*, 321-341.

Cook, T. D. & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings.* Boston: Houghton Mifflin.

d'Appollonia, S. & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*, 1198-1208.

Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunnette & L. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology*, (Vol. 4, 2nd ed., pp. 271-340). Palo Alto, CA: Consulting Psychologists Press.

Glass, G. V., McGaw, B. & Smith, M. L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.

Greenwald, A. G. (1997). Validity concerns and usefulness of student ratings of instruction. *American Psychologist, 52*, 1182-1186.

Greenwald, A. G., & Gillmore, G. M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist, 52*, 1209-1217.

Henninger, M. (2000). Evaluation: Diagnose oder Therapie [Evaluation: Diagnosis or Therapy]? In C. Harteis, H. Heid & S. Kraft (Hrsg.), *Kompendium Weiterbildung - Aspekte und Perspektiven betrieblicher Personal- und Organisationsentwicklung* (S. 249-260). Opladen: Leske + Budrich.

Henninger, M. (2001). Integrative Evaluation: Wie Lehrende unterstützt werden können, Lehre zu verbessern [Integrative evaluation: how teachers can be supported in improving the quality of their teaching]. *Bericht über die Internationale Konferenz für Berufsberatung "Beraten für Bildung, Beruf und Beschäftigung – Neue Herausforderungen" in Berlin, 2000*. Retrievable at http://www.arbeitsamt.de/laa_bb/international/InhaltKongress C/C24henninger.pdf

Henninger,M. (2002, April). *How to avoid teachers' rejection of evaluation feedback*. Paper presented at the 83th Annual Meeting of the American Educational Research Association, New Orleans.

Henninger, M. & Balk, M. (2001). *Integrative Evaluation: Ein Ansatz zur Erhöhung der Akzeptanz von Lehrevaluation an Hochschulen [Integrative evaluation: an approach to enhance the acceptacne of teaching evaluation at universities]* (Forschungsbericht Nr. 133). München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.

Henninger, M. & Balk, M. (2002, September). *Partizipative Evaluation: Ein Ansatz zur Förderung der Akzeptanz von Evaluationsstudien [Participatory evaluation: an approach to foster the acceptance of evaluation studies]*. Vortrag auf dem 43. Kongress der Deutschen Gesellschaft für Psychologie, Berlin, Germany.

Henninger, M., Balk, M., & Mandl, H. (1998). *Doing evaluation in a process-oriented way: Theoretical implications and methodological problems* (Research report No. 95). München: Ludwig-Maximilians-Universität, Lehrstuhl für Empirische Pädagogik und Pädagogische Psychologie.

Henninger, M. & Fuchs, M. (2002, Oktober). *100 Professoren - eine Meinung? Partizipatorisches Qualitätsmanagement und Zielgruppenanalyse an der Hochschule [100 Professors - one opinion? Participatory quality management and analysis of target groups at universities].* Vortrag gehalten auf der 5. Jahrestagung der Deutschen Gesellschaft für Evaluation, Mainz, Germany.

Ilgen, D. R., & Davis, C. A. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology, 49* (3), 550-565.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin, 119* (2), 254-284.

Locke, E. A. (2001). Motivation by goal setting. In R. T. Golembiewski (Ed.), *Handbook of organizational behavior* (2nd ed.). (pp. 43-56). New York, NY: Marcel Dekker.

Locke, E. A. & Latham, G. P. (1990). *A theory of goal setting and task performance.* Englewood Cliffs, NJ: Prentice Hall.

L'Hommedieu, R., Menges, R. J. & Brinko, K. T. (1990). Methodological explanations for the modest effects of feedback from student ratings. *Journal of Educational Psychology, 82*, 232-241.

Marsh, H. W., & Roche, L. A. (1993). The use of students´ evaluations and an individually structured intervention to enhance university teaching effectiveness. *American Educational Research Journal, 30*, 217-251.

Marsh, H. W., & Roche, L. A. (1997). Making students´ evaluations of teaching effectiveness effective. *American Psychologist, 52*, 1187-1197.

Marsh, H. W., & Roche, L. A. (1999). Rely upon SET research. *American Psychologist, 54*, 517-518.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*, 1218-1225.

McKeachie, W. J., Lin, Y.-G., Daugherty, M., Moffett, M., Neigler, C., Nork, J., Walz, M. & Baldwin, R. (1980). Using student ratings and consultation to improve instruction. *British Journal of Educational Psychology, 50*, 168-174.

Miller, M. T. (1971). Instructor attitudes toward, and their use of, student ratings of teachers. *Journal of Educational Psychology, 62*, 235-239.

Murray, H. G. & Smith, T. A. (1989). *Effects of midterm behavioral feedback on end-of-term ratings of instructor effectiveness.* Paper presented at the 70th annual meeting of the American Educational Research Association, San Francisco.

Nease, A. A., Mudgett, B. O., & Quinones, M. A. (1999). Relationships among feedback sign, self-efficacy, and acceptance of performance feedback. *Journal of Applied Psychology, 84* (5), 806-814.

Rindermann, H. (2002). Verbesserung der Lehre. Beurteilungen von Lehrveranstaltungen durch Studierende [Improvement of teaching. Student ratings of courses]. *Forschung und Lehre, 7*, 370-372.

Rindermann, H. & Kohler, J. (2003). Läßt sich die Lehrqualität durch Evaluation und Beratung verbessern? Überprüfung eines Evaluations-Beratungs-Modells [Can teaching quality be improved by evaluation and counseling? Examination of an evaluation-counseling-model]. *Psychologie in Erziehung und Unterricht, 50*, 71-85.

Rotem, A. & Glasman, N. S. (1979). On the effectiveness of students' evaluative feedback to university instructors. *Review of Educational Research, 49*, 497-511.

Spiel, C., Wolf, P. & Popper, V. (2002). Lehre und Lehrevaluation - (un)geliebt? Die Perspektive der Universitätslehrenden [Teaching and teaching evaluation. - (un)liked? The perspective of faculty]. *Zeitschrift für Psychologie, 210* (1), 27-29.

Wilkinson, L. & the Task Force on Statistical Inference. (1999). *Statistical methods in psychology journals*. American Psychologist, 54, 594-604.

Wilson, R. C. (1986). Improving faculty teaching: Effective use of student evaluations and consultations. *Journal of Higher Education, 57*, 198-211.

Young, K. J. & Kline, T. J. B. (1996). Perceived self-efficacy, outcome-efficacy and feedback: Their effects on professors' teaching development motivation. *Canadian Journal of Behavioural Science, 28* (1), 43-51.

Prof. Dr. Dr. Michael Henninger
Institute for Pedagogy I
University of Regensburg
Universitätsstraße 31
D-93040 Regensburg
Tel.: ++49-941-943-3785
Fax ++49-941-943-2450
e-mail: michael.henninger@paedagogik.uni-regensburg.de
www.caimanonline.de

TM034950

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: With or without? The role of reference values in evaluation feedback to teachers	
Author(s): Michael Henninger and Michael Balk	
Corporate Source: University of Regensburg	Publication Date: 04/23/2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

1

Level 1

☑

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.

The sample sticker shown below will be affixed to all Level 2A documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2A

Level 2A

☐

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only

The sample sticker shown below will be affixed to all Level 2B documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

2B

Level 2B

☐

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature:	Printed Name/Position/Title: Prof. Dr. M. Henninger
Organization/Address: University of Regensburg, Universitätsstr. 31, D-93040 Regensburg, Germany	Telephone: 0049 / 941 943 3785 FAX: - 2450
	E-Mail Address: [redacted] Date: 04/30/2003

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC Clearinghouse on Assessment and Evaluation
University of Maryland, College Park
1129 Shriver Lab
College Park, MD 20742